Building a Certification and Inspection Data Infrastructure to Promote Transparent Markets

by Djoko Sayogo

Submission date: 21-Mar-2019 01:32AM (UTC-0700)
Submission ID: 1097127808
File name: nspection_Data_Infrastructure_to_Promote_Transparent_Markets.pdf (751.89K)
Word count: 11702
Character count: 68842

International Journal of Electronic Government Research
Volume 13 • Issue 4 • October-December 2017

Building a Certification and Inspection Data Infrastructure to Promote Transparent Markets

Joanne S. Luciano, Indiana University, Bloomington, USA

Djoko Sayogo, University of Muhammadiyah Malang, Malang, Indonesia & Center for Technology in Government, University at Albany, USA

Weijia Ran, University at Albany, Albany, USA

Nic DePaula, University at Albany, Albany, USA

Holly Jarman, University of Michigan, Ann Arbor, USA

Giri Tayi, University at Albany, Albany, USA

Jing Zhang, Clark University, Worcester, USA

Jana Hrdinova, The Ohio State University, Columbus, USA

Theresa Pardo, Center for Technology in Government, University at Albany, Albany, USA

Deborah Lines Andersen, University at Albany, Albany, USA

David F. Andersen, University at Albany, Albany, USA

Luis Felipe Luna-Reyes, University at Albany, Albany, USA & Universidad de las Americas Puebla, Mexico

ABSTRACT

This article reports on data architecture that reduces information asymmetries to support public-private collaboration to govern product certification and inspection for promoting transparent markets and building consumer trust. The data architecture is a proof-of-concept set of data standards called the Certification and Inspection Data Infrastructure Building Block (CIDIBB) for data storage, retrieval, sharing and automated reasoning of data that can be used to respond the question: what constitutes a trustworthy certification and inspection process? CIDIBB consists of three interrelated ontologies, focusing specifically on certified fair-trade coffee that has the potential to become universally applicable to any certification and inspection process for products or services. The evaluation results suggest that CIDIBB is able to test the trustworthiness of certification schemes, providing consistent results. CIDIBB will contribute to support public-private collaboration to solve public problems such as the promotion of sustainable production and fair labor practices.

KEYWORDS

Building Block, Certification, Data Infrastructure, Inspection, Ontology, Sustainable Production, Virtual Certificates

DOI: 10.4018/IJEGR.2017100104

1. INTRODUCTION

Classic market economic theory of supply and demand works under the assumption of perfect information—both sellers and buyers have access to full information about the state of the market (Stigler, 1957). Unfortunately, these assumptions about information in free markets are often not true; information asymmetry clouds the relationship between buyers and sellers in the market. When we buy a pair of running shoes or a pound of coffee, for example, we do not know if they were manufactured using child labor, exploiting workers or damaging the environment. Governments, NGOs and private organizations have developed strategies to reduce information asymmetries such as labeling and certification, chain-of-custody, and infomediary platforms. The third-party certification and labeling industry, for instance, has expanded rapidly since the 1990s (Albersmeier, Schulze, Jahn, & Spiller, 2009; Jahn, Schramm, & Spiller, 2005). Private organizations have also increasingly campaigned for chain-of-custody – the ability to trace the path of products from producers to consumers.[1] Both third party certification and chain-of-custody rely on labels attached to the product, thus on the clarity and verifiability of such labels (Starobin & Weinthal, 2010). Consumers, unfortunately, often do not have the ability to drill down the information behind the label in order to make informed choices. Moreover, the rapid proliferation of labeling further obstructs the ability of consumers to understand the meaning behind labels, making them no longer adequate to provide warranty of trusted information (Jarman & Luna-Reyes, 2016).

Furthermore, harsh competition for contracts has rendered third party certifiers with the high risk of false incentives and adverse selection (Albersmeier et al, 2009), and many third-party certifiers are lacking credibility and their schemes are proven fallible by the increasing number of fallacious and vastly exaggerated claims (Starobin and Weinthal, 2010). Arguably, governing such a complex market requires collaboration between government and private entities. Unfortunately, standardized data, tools, and applications that could facilitate sharing information to support efficient collaboration between government and private entities is yet to exist. For that reason, in this paper, we propose that the provision of a data architecture and data standardization potentially could alleviate the difficulties in developing a public-private collaborative governance for promoting transparent markets. This architecture constitutes an ontology-based building block for a system that enables standardized reporting of certification and labelling practices, including the potential for supporting the identification of a trustworthy virtual certificate. That is to say, this paper presents a set of ontologies and an assessment framework that can be used to respond to the question: what constitutes a trustworthy certification and inspection process?

We introduce the concept of a Certification and Inspection Data Infrastructure Building Block (CIDIBB), combining the ontologies and a process that involves the use of 28 questions to assess the trustworthiness of any given certification. The assessment results demonstrate the indispensable function of governance mechanisms to make available the necessary information to reduce information asymmetry in the market, which is a significant contribution of the paper. The development of the 28 questions as well as the set of ontologies were specified, conceptualized, implemented, and evaluated based on data collected through interviews, the focus group, the survey, and archives. Technology infrastructures constitute only one of the components for realizing a more transparent market for consumer products (Graham and Haarstad, 2011). The research reported here, nonetheless, is part of a larger project called I-Choose that focuses on building information sharing networks to support consumer choices. The project includes both the technology components in terms of data standards and procedures (CIDIBB) as well as governance and policy components (Jarman & Luna-Reyes, 2016).

The paper is organized in 7 sections including this introduction. Section 2 summarizes previous research in FIPP systems and ontologies. Section 3 describes the general approach in building and testing CIDIBB. Section 4 includes a brief description of the main components of CIDIBB as a set of three ontologies, CertIN, FLO, and CiTruST. Section 5 presents an empirical evaluation of CIDIBB, showing ways in which CiTruST can be used to automatically classify certification systems in terms of their trustworthiness. Section 6 includes two potential scenarios as examples of the use of CIDIBB in practice. Finally, section 7 includes concluding remarks and future work to fully develop CIDIBB.

2. PREVIOUS RESEARCH

In this section, we introduce the concepts of Virtual Certificates and Full Information Product Pricing (FIPP) systems. The section also discusses the importance of trust in FIPP systems and summarizes previous research in FIPP systems and ontologies.

2.1. Physical Versus Virtual Certificates

For over a century, various forms of certificates have been attached to products on retail shelves or in product advertisement. Certificates can be sponsored by public, private and non-for-profit organizations. The Good Housekeeping Seal, for example, has been a well-known marker of a product's tested and certified quality since 1909. Other examples include the Fairtrade certificate and the USDA Organic seal, which certify specific processes followed in the production process. The US Department of Energy's Energy STAR rating is a physical certificate that provides information on an appliance's annual energy consumption as well as a comparison of that specific appliance with other models in its class. Every new car sold in the United States has attached to it a certificate that declares its fuel economy using methods approved and governed by the US Environmental Protection Agency.

With the advent of the Internet, new virtual certificates have been created. These virtual certificates use a unique identifier to link an information package to a given product or service, offering more detailed information than physical certificates, usually accessible through the Internet. Some of these virtual certificates are online versions of certification and rating activities that had formerly been the basis for physical certificates. (See, for example, www.goodhousekeeping.com/product-reviews/seal/, www.energystar.gov, or www.fueleconomy.gov). Consumer Reports has transformed its traditional magazine-bound rating system into an online subscription service that provides head-to-head ratings of many products in the same market niche (www.consumerreports.org/cro/index.htm).

Other virtual certificates move well beyond the automation of existing physical certificates. For example, GoodGuide (www.goodguide.com) uses a summary 10-point scale to rate a wide range of personal care, food, household, and other products aimed for babies and children on the triple dimensions of their health, environmental, and social impacts. Online virtual certification systems provide a rich and varied package of information, moving far beyond what can be communicated with a simple physical certificate describing select consumer products and services.

To be useful, both physical and virtual certificates need to be trustworthy. The recent scandal involving the VW Corporation's reporting of fuel efficiency data for its diesel fleet underscores the fact that even government-sponsored certificates need to be carefully supervised and scrutinized for accuracy (Davenport & Hakim, 2016). Even in the absence of scandal, confusion may grow about what a certificate actually means or how compliance with standards is measured. For example, while most consumers recognize the Fairtrade label, a much smaller pool of consumers could actually describe its meaning or distinguish among different Fair Trade certificates.

2.2. Ontologies, the Semantic Web and Virtual Certificates

In this work we propose that current Semantic Web Technologies constitute valuable tools to develop trusted virtual certificates. In the field of information and computer sciences, ontologies refer to explicit specifications of terms and their relationships within a domain of interest (Gruber, 1993). Such specifications provide a number of benefits, the most basic of which is enabling a computer to reason over the terms and properties of data (Uschold & Gruninger, 1996). Semantic web applications or services require that data be published in a format that makes use of the specifications established in the ontologies" (Berners-Lee, Hendler... etc.). Data published following such specifications may be called "linked data," and such data serves as building blocks for the semantic web (Berners-Lee, 2006). In this work we adopt these semantic technologies for establishing specific ontologies of the certification and inspection process. These technologies thus make up CIDIBB, establishing specific data format and the web platform for publishing information. Creating data this way allows for more precise results from searches in the web and the automation of inferences over contents of the data (Bizer, Heath, & Berners-Lee, 2009). In more technical terms, using ontologies for the semantic web involves publishing data in the Resource Description Framework (RDF) file structure (W3C specification) in which subjects, predicates and objects (or RDF triples) within components of the data are explicitly identified.

As semantic web technologies make use of specifications established in domain ontologies, they make it possible for data from different organizations and with different terminology (e.g. certification and inspection processes) to be integrated and classified in a structured way to improve search and enable the use of automated reasoning. For example, when a certification or inspection organization provides data where the "field inspector" is labeled as an "auditor," since they refer to the same entity or function, definitions in the ontology may indicate that these terms refer to the same concept. A software application can then use the ontology to determine that two attributes in two different datasets are equivalent. Applications can also use inference tools to make determinations about items and properties included in the data set, such as: "Is there an auditor?" or "Is the date of inspection before the date of certificate?"

Nowadays, although XML-based semantic technologies have been widely used to create standards for data sharing and exchange in various domains – e.g., eXtensible Business Reporting Language (XBRL), a XML-based standard to present financial data and allows data exchange and integration (Henderson, Sheetz, and Trinklec, 2012), or LandXML, a XML-based standard to facilitate the exchange of data related to civil engineering processes (LandXML, 2017) – few ontology-based semantic technologies are adopted. In the domain of certification and inspection processes, no ontology-based semantic technology data standards have been found in literature, let alone applications to facilitate information sharing and use, based on those standards. However, semantic web technologies provide a robust and scalable environment for the development of virtual certificates and the data for which they are needed. Therefore, ontology-based technologies are used in this research as the framework for efforts to design, build and test the concept of a Certification and Inspection Data Infrastructure Building Block (CIDIBB) that supports data exchange, integration, and automatic reasoning.

2.3. FIPP Systems and Trust

Most product certificates involve a network of organizations promoting what we have called a FIPP system (Luna-Reyes et al., 2014). The approach to producing FIPP systems involves the creation of a certification ecosystem. Certifying organizations will set and make public standards for all types of products and services. Third-party certifiers will inspect facilities, processes, and outcomes to certify that they indeed meet the standards. Finally, a (hopefully) trusted certificate will be attached to the product or service to provide consumers with the information that they need; i.e. perfect information without asymmetric bias. Consumers will pay a premium in exchange of the additional product information attached to the certificate.

Previous research has also shown that trust plays a key role in all FIPP relationships (Luna-Reyes et al., 2013; Zhang et al., 2016). In fact, trust is considered as an alternative governance mechanism in most collaborative relations (Powell, 1996; Puranam & Vanneste, 2009). Moreover, the literature points out the importance of trust in these market transactions, particularly in the case of unobservable product attributes (Arora, 2006). The general understandings of the nature of trust in the literature include vulnerability, risk, and the role of positive expectations or optimistic belief (Rousseau et al, 1998).

Researchers have identified several mechanisms for "trust production," which include calculative, relational and institutional mechanisms (Rousseau et al, 1998). Institution-based trust refers to the existence of an institutional framework that regulates the relationship between trustee and trustor (McKnight & Chervany, 2001; Rousseau et al, 1998). The favorable conditions enabled by the existence of institutions and regulations provide assurance of security and control over risks which in turn induce trust (McKnight & Chervany, 2001). The fact that institution-based trust strengthens interpersonal trust is very relevant for systems such as the proposed CIDIBB (McKnight & Chervany, 2001). Given the complexity of supply chains which result in limited interpersonal trust (Campbell, Murcott & Mackenzie, 2011; Starobin & Weinthal, 2010), consumers increasingly rely on institutional trust such as the third-party certificates. Finally, from an economic perspective, trust can be developed based on a calculated analysis of risks versus benefits (Williamson, 1993; Pavlou & Gefen, 2004).

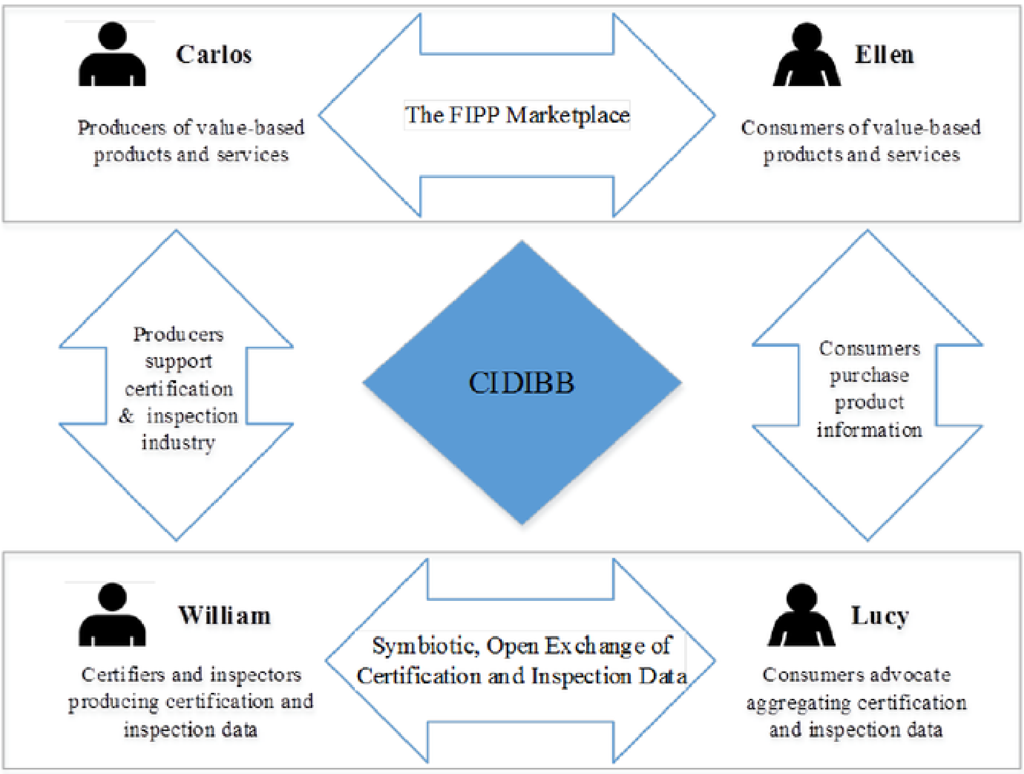
Research has also posed that trust development is related to the trustworthiness of the trustee (Colquitt, Scott & LePine, 2007), and the trustor's propensity to trust (Mayer, Davis & Schoorman, 1995). Trustworthiness represents the character and competence that inspire positive expectation on the trustee (Colquitt et al, 2007). Research on trust agrees that the bases of trustworthiness consist of three components: benevolence, integrity and ability (Mayer et al, 1995). Assessing whether the trusted party can benefit from being trustworthy based on rationality depends on reliable information (Hart & Saunders, 1997; Nidumolu, 1989; Wang & Benbasat, 2005) and openness about sharing information (Luna-Reyes et al, 2013).

The challenge of making trustworthy virtual certificates a reality lies mainly then in making the vast amounts of disparate data shareable and understandable across certification and inspection processes in a way that will be trusted by consumers. A needed component is a combination of data standards and procedures that allow data to be shared seamlessly among the potential users of those data. This component is referred to as the "Certification and Inspection Data Infrastructure Building Block (CIDIBB)."[2] CIDIBB is a set of data standards in the form of a formal ontology of the certification and inspection process that would allow the creation of a data ecosystem for certification processes.

Figure 1 shows the main components of the information ecosystem created around the CIDIBB. Possible stakeholders are simplified into four representative characters —Ellen, Carlos, Lucy and William. Ellen, shown in the upper right-hand quadrant, represents consumers who are seeking information from a virtual certificate to make a purchasing decision (imagine coffee that is sustainably produced). Ultimately that information must originate from a coffee producer like Carlos (think of Carlos as a Mexican cooperative disclosing information about its coffee farming practices). However, Ellen cannot directly interact with Carlos as she makes her purchase, so she gets information from Lucy. Lucy represents a consumer advocate industry (like Consumer Reports or GoodGuide) that analyzes the full information package of consumer products and then provides that information to consumers such as Ellen. Lucy will use the CIDIBB framework to ascertain the trustworthiness of a (virtual) certificate. Lucy relies on William, a member of the inspection and certification industry (for example inspectors in Mexico visiting Carlos' cooperative), who uses CIDIBB to broadcast information about how, when, where, and by whom consumer products are created. William will use CIDIBB ontologies to create a standardized report of Carlos practices as they relate to a specific certification. Finally, Carlos, the producer who seeks to sell his products to Ellen, is cooperating with William to certify his production processes and to document unobservable attributes of his products because he understands that Ellen is willing to pay a price premium for products produced using methods that are congruent with her values.

The marketplace will drive the content of virtual certificates. For example, if consumers are concerned about the environmental impacts of the products they buy, then William's virtual certificates would focus on, for example, the carbon footprint created in producing and delivering the product to the final consumer. However, the system only works as long as Ellen continues to trust the information about virtual certificates that are being introduced into this newly formed full information product pricing (FIPP) information ecosystem.

Figure 1. Certification and Inspection Data Infrastructure Building Block (CIDIBB)



3. RESEARCH DESIGN AND METHODS

The research reported herein results from a three-year project to better understand the barriers and enablers of support supply-chain interoperability in an effort to provide trusted information about products and services to consumers. The research team consists of a network of researchers and practitioners from Canada, Mexico, and the United States, studying the single case of coffee grown in Mexico, and distributed and consumed in Canada and the United States.

The project involved the 5 sequential steps introduced in Table 1. We used interviews, a focus group, and archives to understand the knowledge domain of certification and inspection processes. We then extracted the concepts and relationships among concepts that specify and conceptualize the structure and processes of the certification and inspection domain. The extracted concepts and relationships were then used to construct classes, relations, attributes, and instances of proposed ontologies. We created scenarios and competency questions from the data for the evaluation of implemented ontologies. Initial steps in the process lead to the primary question to be asked to the system: what constitutes a trustworthy certification and inspection process? Third-party certification and inspection processes were subsequently mapped to develop three ontologies, using commonly used ontology-development methods (Gruninger & Fox, 1995; Fox & McGuinness, 2008; Uschold & Gruninger, 1996). Details on ontology development and research methods used in the project can be found elsewhere in the literature (Sayogo et al., 2016; Jarman & Luna-Reyes, 2016).

This paper reports on the last step in the process, the development and use of 28 questions as a framework to evaluate the utility of the ontologies as instruments to assess the overall trustworthiness of certification and inspection data (see Table 2). The development of the questions was informed by all previous project developments, and were used as a normative definition of trustworthiness. In

Table 1. Research methods and empirical evidence

Step	Description	Activities	Data Collection
1	Developing the use case	Exemplar of fairtrade coffee procurement	Interviews with (for exemplar): Roaster Rue Champagneur, Canada Coop – Tosepan Titataniske Mexico
		Focus group discussions with network members	
		Interviews with producers, roasters and certifiers	
		Document analysis	
2	Map the structure of certification and inspection data	Interviews with certification body (Control Union, Fairtrade USA and Fair for Life)	Composition of Interviewees: 9 Producers & exporters 5 Roasters & importers 5 third party certification 6 NGOs
		Archival analysis examining documentation of Flo-Cert	
		Mining data content from exemplar and Audit Report	
3	Developing ontology of certification	Interviews to refine the focus	Focus group (23 participants) 13 from academics 1 from certifying agencies 5 from consumers advocate 1 from state government 2 from NGOs 1 from retailer
		Document analysis to identify semantic components	
4	Converting tabular data to triple data	Open source conversion tool csv2rdf4lod to convert the data format	Survey questionnaire: 159 respondents from USA and Mexico
		Openlink Virtuoso to convert into triple data	
5	Analysis of the 28 use case questions	Develop 28 trust questions Distributing survey questionnaire	Document analysis of 3rd party certification
		SPARQL query	
		Inference-based retrieval of data	

practice, the 28 questions will be used primarily by power users (such as Lucy, the consumer advocate) as a way to query and aggregate information depicting trustworthy certification schemes (Section 6 introduces a more detailed scenario).

Consistent with previous research on trust (Hart & Saunders, 1997; Nidumolu, 1989; Wang & Benbasat, 2005; Luna-Reyes et al., 2013), the trustworthiness evaluation system is based on two values, rightfulness and transparency. As such, the certification trustworthiness evaluation system consists of two major components: a) trustworthiness evaluation criteria and b) data openness indicators. Ideal trustworthiness was assigned when all evaluation criteria were met. Inability to fulfill the criteria decreases the trustworthiness level. The degree of data openness and exchange consists of two competing factors, data availability (data source) and governance level needed to extract the data. The less transparent the data is, the higher the governance level is needed to extract the data and vice versa. If the criteria cannot be met using these two indicators, it will decrease the trustworthiness level of the certification scheme. The measurement for these two indicators is listed in Tables 3 and 4. Data Source (availability) measurement consists of five possible sources of data, as shown in Table 3. There are three different levels of governance with specifications listed in Table 4. A more complete discussion of how each of the 28 questions were classified according to the three categories in Table 4 can be found at https://github.com/jluciano/ichoose.

Table 2. Certification trustworthiness evaluations criteria (28 use case questions)

No.	Evaluation Criteria	Assessment Indicator	
		Data Source (A ↔ F)	Governance (1 ↔ 3)
1	Is certification standard openly published (available on a website)		
2	Is certification compliance criteria/control points openly published (available on a website)		
3	Can know date of the inspection		
4	Can know date of certification		
5	Can know who is the inspector/auditor		
6	Can know how non-conformities are handled by the applicant		
7	Can know who is the standard-setting body		
8	Can know what type of organization made the standard (government, private, Non-profit)		
9	Can know who gives accreditation to the certifier		
10	Can know when the standard setting body was established		
11	Is inspection report signed by an inspector		
12	Is certificate signed by the certifier		
13	Can know location of audit/inspection		
14	Is the list of non-conformity (measured score below standard) information available		
15	Can know the accreditation body of the standard-setting organization		
16	Can know the certification bodies of a particular standard-setting organization		
17	Can know the certification body of a particular applicant for particular products		
18	Does inspector/auditor have license		
19	Inspection/audit results openly published (available on request by FOIA or NGO)		
20	Does certification standard conform to a government-backed standard, e.g., USDA, EU-ECO-Regulation		
21	Does certification standard conform to standard within an inter-governmental organization (e.g. ILO)		
22	Who sponsor the development of the standard (consumer NGO, producer, manufacture)		
23	Can know who translated the standard into compliance criteria/control points		
24	Is standard-setting body independent from the accreditation body, such as ISO		
25	Is certifier independent from the accreditation body		
26	Is certifier independent from the standard-setting body		
27	Is inspector/auditor independent from the standard setting body		
28	Is inspector/auditor independent from certifier		

Table 3. Data availability (source) levels

A	If data is available by searching the Web.
B	If data is available in the certification and inspection database. This is the database available in the certifier system specifically to store information and data related to certification and inspection results.
C	If data is available in regulator's database. The regulator here could be government agencies such as USDA (United State Department of Agriculture) or Self-Regulated Organization such as ISO (International Standard Organization).
D	If data is available in the information system of the certifier but not in the certification and inspection database. An example of this database is the human resource database; detail information of inspector is only available the HR database of certifier.
E	If data is available in the database of a standard-setting organization. In a majority of certification schemes, the certifier is independent of a standard setting organization.
F	If the source of this data is not explicit and cannot be easily located.

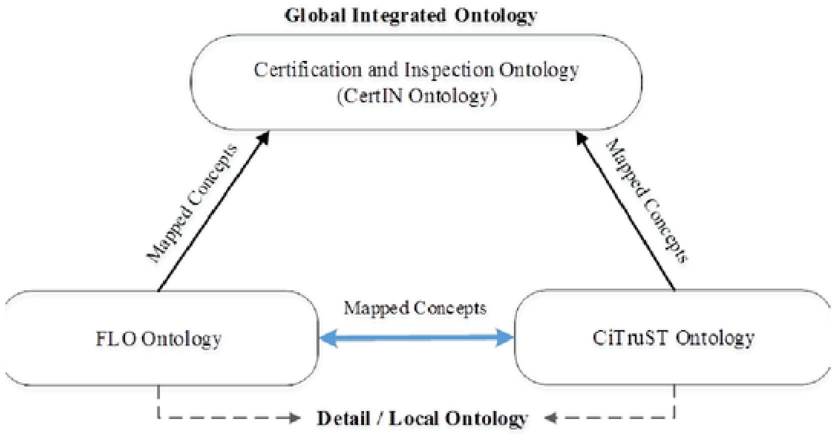
Table 4. Governance levels

1	There is no need to appeal to higher governance authority to access the data. This assumes that the data is available in the certification and inspection database and the certifier agrees to release the data.
2	May need to appeal to higher governance authority to access the data because this data might exist in multiple data sources and one of the sources is outside of the certifiers' and standard-setting organization's information system.
3	Higher governance intervention is needed to be able to answer the question because the source of the data is not explicit or cannot be easily identified.

4. KEY TECHNICAL COMPONENTS OF CIDIBB: ONTOLOGY BASED DATA STANDARDS AND EVALUATION SYSTEM

In this section, we briefly describe the main components of CIDIBB as a set of three ontologies, CertIN, FLO, and CiTruST (see Figure 2). These three ontologies together form the fundamental base of the proposed CIDIBB, an abstract architecture for data storage, retrieval and automated reasoning of certification and inspection data. CertIN Ontology defines the high-level abstraction of concepts, which we refer to as the upper ontology. FLO Ontology and CiTruST we refer to as

Figure 2. The proposed ontologies and their relationships (adapted from Sayogo et al., 2016)

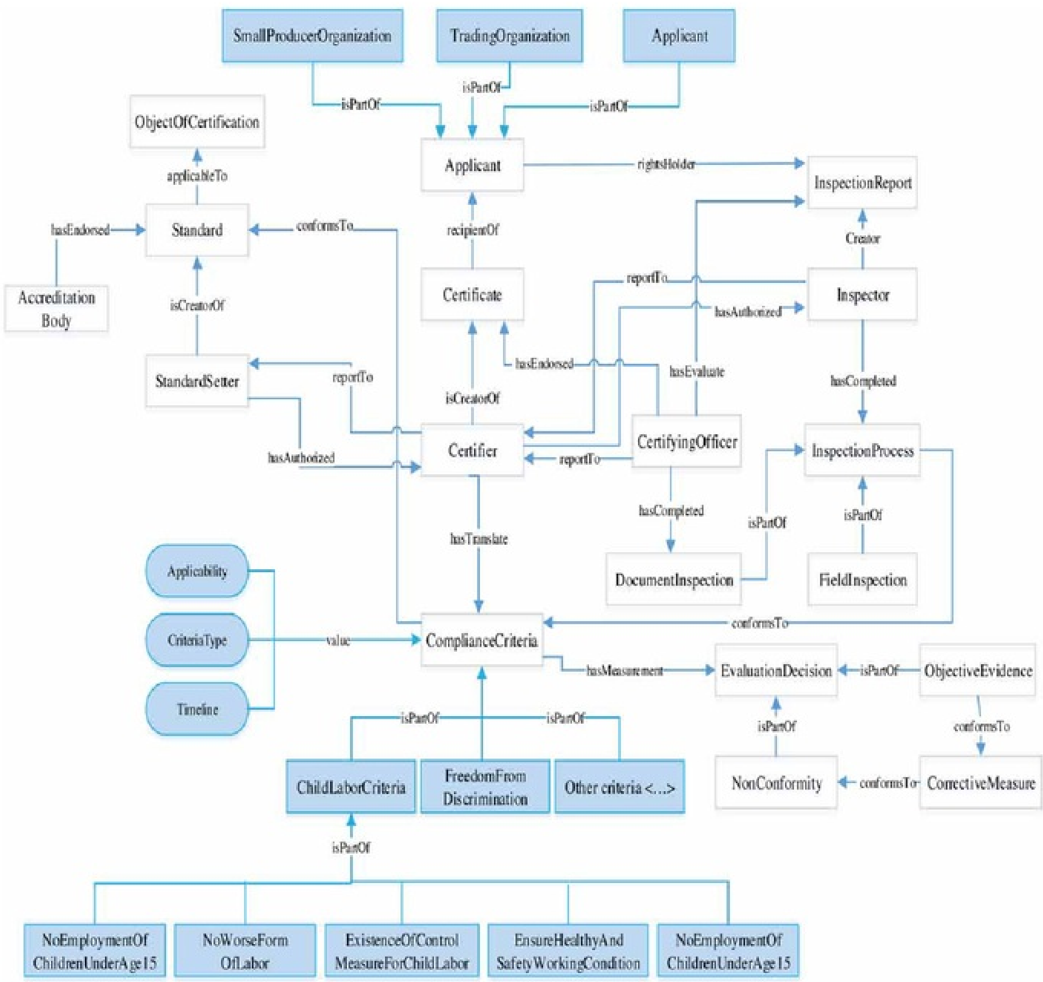


the local ontologies. They inherit and expand high-level concepts defined in the global ontology, as shown in Figure 3. Following (Noy, 2005), the CertIN Ontology serves as an interlingua through which the ontology-to-ontology mapping was conducted. The rest of this section briefly describes each ontology. For more elaborate description about these ontologies and their development process, see please refer to Sayogo et al (2016).

CertIN provides the higher-level definition of a certification system that serves as an overarching architecture to connect multiple, more detailed ontologies for each certification and labeling scheme. To ensure compatibility and interoperability, the development of the CertIN ontology used standard definitions of class and property that are available from existing ontology literature. In addition, CertIN has adopted classes and properties from three ontologies recommended by the W3C (World Wide Web Consortium). These three include: Dublin Core (http://dublincore.org/2008/01/14/dcterms.rdf), FoaF (http://xmlns.com/foaf/spec/) and Good Relation (http://www.heppnetz.de/projects/goodrelations/) (Sayogo et al., 2016).

There are five major components in CertIN: agent, document, object of certification, inspection process in certification, and evaluation decision. Agent refers to entities, either individual or

Figure 3. Connecting CertIN and FLO ontology (adapted from Sayogo et al., 2016)



Notes: Blue rectangle or rounded rectangle represent FLO Ontology organization, that have different roles in a certification system (applicant, certifier, certifying officer, inspector, and standard setter). Document represents three major documents that are used or produced as an outcome of certification (certificate, inspection report, and standards and compliance criteria). Object of certification refers to three objects that are the focus for certification (products, processes, and business entities). Inspection process—document inspection and field inspection—represents the process of gathering evidence to assess the compliance of an applicant or object of certification with the standard and compliance criteria set by the certifier and standard body. Evaluation decision refers to three decisions—non-conformity, corrective measure, and objective evidence—that represent the applicant's conformance to the certification standard and criteria. The concepts (classes) and and their relationships (properties) in CertIN are presented in Figure 3. Rectangles represent classes, and arrows represent properties. A property defines the relationship between two classes.

CiTruST ontology uses the classes and properties from CertIN to define the quality of a certification process. We started with the basic structure of a certificate to find indicators for the quality of certification. Some components of the basic structure of a reliable certification process are accreditation body, certification body, standard setter, and monitoring process (Albersmeier, Schulze, Jahn, & Spiller, 2009; Deaton, 2004; Jahn, Schramm, & Spiller, 2005; Tanner, 2000). The document analysis and interviews further indicated the importance of independence and monitoring processes that combine both document and field inspection as an indicator of reliable certification. The existence or nonexistence of particular components in the structure of certification indicates the degree of reliability of the certification scheme. For example, an independent certification body provides more reliable monitoring than an internal certification body due to the elimination of conflict of interest (Deaton, 2004).

Following this logic, two major classes for CiTruST are proposed: Object of trustworthiness and level of trustworthiness. Object of trustworthiness refers to information from which users can draw inferences on whether a particular certification is trustworthy. Object of trustworthiness encompass all classes in CertIN ontology: agent, documents, evaluation decision, and inspection process. Level of trustworthiness refers to the degree of certification trustworthiness derived from the conformance or non-conformance to the object of trustworthiness. The object of trustworthiness refers to the classes specified in CertIN ontology minus object of certification. CiTruST ontology proposes four levels of certification process reliability, A to D. The assignment of the level depends on the existence of the criteria in the object of trustworthiness. The criteria/classes to measure the level of trustworthiness in the CiTruST Ontology are based on the certification trustworthiness evaluation system.

The FLO ontology is an example of a local ontology created from Fairtrade certification and inspection processes to further demonstrate how the CertIN ontology can be mapped to specific certification and inspection schemes. The ability of CertIN to map into a local ontology such as FLO enables users to extract consistent and detailed information for assessing the trustworthiness of a certification scheme. The classes in the FLO ontology represent detailed sub-classes in the CertIN ontology. For instance, class Applicant in CertIN is classified further into different sub-classes that pertain to FLO certification: SmallProducerOrganization and TradingOrganization. Each class in the FLO ontology will have further detailed properties that comply with FLO requirements.

The most important elements of the FLO ontology are the detailed classifications of compliance criteria into their properties. A compliance criterion is constructed with several restrictions, as defined in the FLO standard, by specific timeline, criteria types, measurement of the criteria and organization applicability. These restrictions represent the properties of the criterion. Conformance to these properties affects the evaluation decision for certification and it is also argued that conformance to these properties defines the level of trustworthiness of the certification schemes.

5. EMPIRICAL TESTING OF THE PROOF OF CONCEPT

This section describes the empirical evaluation of CIDIBB, showing ways in which CiTruST can be used to automatically classify certification systems in terms of their trustworthiness.

5.1. The Process of Examining CIDIBB

5.1.1. Generate a Fair Trade Sample Data Set

To test how CIDIBB could work, the structure of Fair Trade certification and inspection data was mapped. Data tables were created to represent: a) a certification body database structure that supports certification and inspection, and b) data aggregated by an information aggregator. Data tables representing certification data were classified as the Certification and Inspection Database (CID) to refer to a database that consists only of certification and inspection results. Data and documentation is available in our open access repository, https://github.com/jluciano/ichoose.

A total of eight data tables were created to represent the data structure of Flo-Cert, the main certification body under FLO. The data tables are: contact data, product code, certification status, audit status, audit results, audit workflow status, corrective measure and objective evidence, and inspection checklist data. These eight data tables consist of 81 data attributes with each table comprising eight to 13 attributes. Finally, two data tables with 14 data attributes were created to represent the list of data collected by an information aggregator. Two synthetic certification bodies named “Dave and Nic” certification body and “Non-Violent Dove” certification body were also created. These two synthetic certification bodies were designed to follow practices less stringent than FLO practices and broadly congruent with “light green-washing” (Dave and Nic) and “heavy green-washing” (Non Violent Dove).[3]

In addition to the database described above, Data tables were manually created to represent the data structure that might be created by an information aggregator. Information aggregators in this ecosystem search the Web to extract data, and then they refine and re-format the data for easy use. This dataset represents data about certification and standard bodies that are not within the more narrowly defined CID containing only information about direct certification activities. Hence, information outside the CID contains data such as data tables with information about the governance and history of different certification and standard bodies.

5.1.2. Publish Data Set as a RDF Triple Store

The data in the CID was formatted as standard tables of data such as those that might be found in a spreadsheet (such as EXCEL) or a relational database (such as ACCESS). Using the classes and relationships defined by the CertIN and FLO ontologies, standard semantic web technologies were used to recast those tables of data as a RDF triple store that is searchable using SPARQL queries. An example of converted RDF file is as follows:

flo-certification:FLO_1341
rdf:type certin:Applicant;
dcterms:identifier “FLO_1341”
certin:hasLocation “Mexico”
certin:hasName “Cooperative Coffees Inc
certin:serviceType “Trader-Payer”
certin:hasCertificationOfficer “Janssen Martina”
certin:hasInspectionReport <http://ichoose.tw.rpi.edu/source/ctg-albany-edu/dataset/
flo-certification/Inspection/AO-0045>
certin:hasCertifier Flo-Cert
ov:csvRow “2”^^xsd:integer

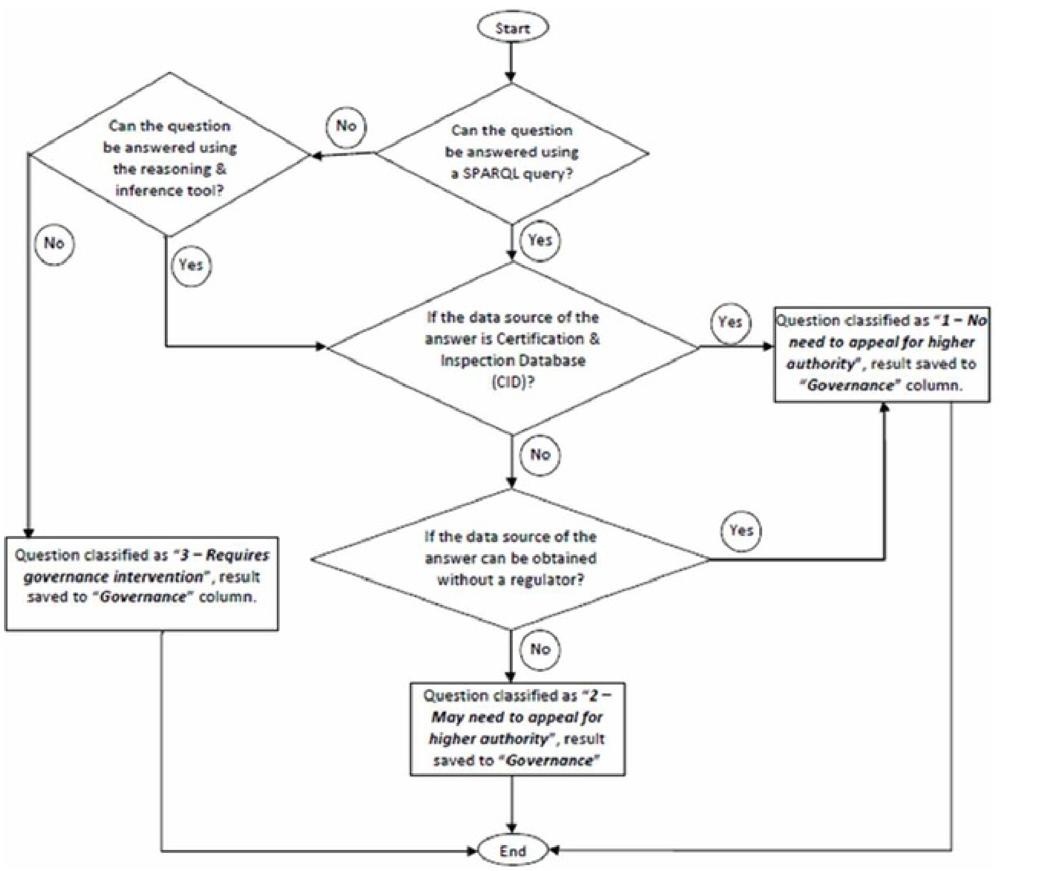
5.1.3. Run SPARQL Queries Against the Data

Testing the proof-of-concept begins by running a SPARQL query against the data in the triple store to see if the basic questions could be answered by such a direct query (see Figure 4). If the answer to the question could be retrieved using a SPARQL query or could be answered using the reasoning and inference tool, then the data source of the answer could be examined. Questions whose answers can be found directly in publicly available data or inferred by data provided by the certifying and inspection agencies without any need to appeal to a higher authority were classified as "Level 1: No need to appeal for higher authority." For example, "Can I know the date of inspection?"

A second class of questions can still be answered either by a SPARQL query or by using the reasoning and inference tool, but the answer to these questions does not originate with the certifying and inspecting processes and organizations per se. These questions seek information about the certification and inspection organizations and processes themselves. Hence, answers to these questions require that data be made available in the triple store that refers to some higher authority. These questions were classified as "Level 2—May need appeal to a higher authority." For example, the question "Can I know the accrediting agency for the standard setting body?" appeals to a higher authority to provide the name of the accrediting agency as this data is not available in the CID.

Finally, if questions could not be answered by a direct SPARQL query nor could be inferred using formal inference tools or manual curations of the data, these questions were classified as "Level 3—Requires Governance Intervention." For questions in this category, the data sources of their answers are not explicit, and governance interventions are required to locate the answer and make it publicly

Figure 4. The process of examining CIDIBB usefulness in trustworthiness evaluation using 28 questions



available. For example, the question "Is the inspector/auditor independent from the standard setting body?" cannot be answered without intervention by some form of governance. Finally, the number of Use Case Questions could be answered at each level was summarized.

In this way, a question can be answered using a SPARQL query in the first two cases, when the data is available in the CID, and when the data is available in other data points published online. For example, "Q7, Can we know who the inspector is?" The inspector's name was saved in the CID, so this question can be answered by running a simple SPARQL query against the CID-converted triple store, and inferences are not required. As for Q19, "Can we know when the standard setting body was established?" The date of establishment of the standard setting body is not available in the CID, so this question cannot be answered by running a SPARQL query against the CID-converted triple store. However, if the standard setting organization or other regulatory agencies provide this information freely and make it available as data points on the web, this question can be answered by running a SPARQL query against the triple store linked to these data points. In comparison with Q7 and Q19, some questions cannot be answered by running a simple SPARQL query, for example, "Q26, Is the inspector independent from the certifier?" The data source of the answer to this question is not available in the CID-converted triple store. Also, answering Q26 will need logical inference from facts and evidence, such as the policy on conflicts of interest for the inspector. The high level of governance is needed to ensure the "independency."

5.2. Summary of Results

The testing process described in the previous section was applied to the datasets stored using the CIDIBB architecture. Answers to the 28 use case questions for each of the four datasets produced a unique distribution across the three classified levels (See Figure 5). Differences in the level of difficulty required us to retrieve the answers to these 28 questions can be used to assess the relative trustworthiness of various certification and inspection processes. The results clearly distinguish between high quality FLO data and data from the other virtual certificates that were missing answers to many of the detailed questions in the use case (See Figure 5). An "Ideal Benchmark" certificate was added to characterize a hypothetical virtual certificate that could answer 100% of the questions posed by the use case.

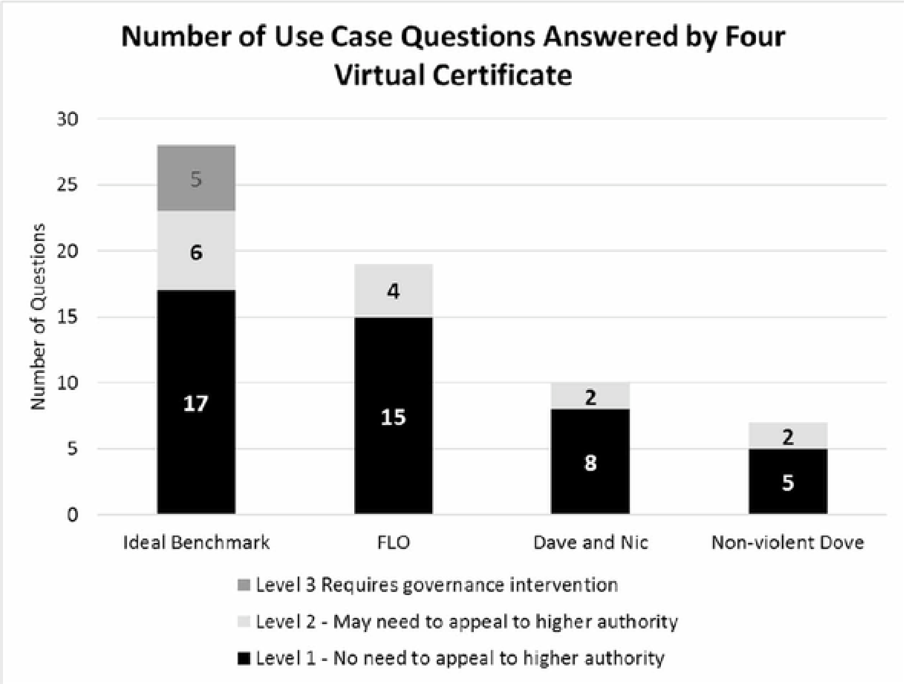
Tautologically, the Ideal Benchmark provides answers to all 28 questions in the use case whereas the FLO certificate answers 19 of the questions; the lightly green-washed certificate ("Nic and Dave") answers ten of the questions, and the heavily green-washed certificate answers only seven of the detailed questions in the use case. Green-washed systems cannot "hide" the fact that their certificates are based on short cuts and less than rigorous methods. Especially noticeable is the sharp decline in questions that can be answered directly by SPARQL queries. By testing the criteria using both SPARQL and DL queries, it demonstrates that not only is CIDIBB able to test the trustworthiness of certification schemes but also that the ontology generates consistent results. It is important to notice that even in the ideal benchmark, governance intervention is required in order to respond at least 5 out of the 28 questions. This result suggests that governance mechanisms play an indispensable function in making available the necessary information to reduce information asymmetry in the market.

In this way, the CIDIBB architecture can support a system to integrate and exchange information, allowing consumer advocates, such as Lucy, to directly query such data for answers to the 28 use case questions and to use those answers to inform consumers, such as Ellen, about the trustworthiness of the certificates on the products and services she plans to purchase.

5.3. Using CiTruST to Automatically Classify Trustworthiness of Certification Schemes

As discussed in the previous section, a skilled human user of the CIDIBB can exhaustively query the existing data for multiple certification schemes, paying close attention to all 28 use case questions to arrive at the results presented in Figure 5. However, that process would certainly be time-consuming

Figure 5. The result of empirical testing of the certification schemes into the CIDIBB benchmark for trustworthiness



and error-prone. Fortunately, using the reasoning functions on ontology-based data, the task of assessing trustworthiness can be automated using the CiTruST ontology, which structures the task of assessing global trustworthiness *as normatively defined by 28 use case questions*. The manual process described in the previous sections are automated to classify a certification scheme as of four types (A through D) where an "A" classification is compatible with highly trusted data (again as defined by the 28 use case questions) and "D" classification is compatible with heavily green-washed certification processes. Table 5 presents the results of the automatic classification of trustworthiness which compares well to hand-calculated results shown in Figure 5.

In the test, the existing structure of the dispersed data associated with the FLO certification and inspection process for coffee was carefully documented. Next, queries structured by the CIDIBB were used in two ways:

1. SPARQL queries coupled with inferences as described in section 3 for all three virtual certificates;
2. DL queries plugin in Protégé by querying the ontology using the 28 trustworthiness use case questions as competency questions.

Table 5. The automated trustworthiness ranking of three certification schemes using the CiTruST ontology and reasoning

No.	Certification Scheme	Trustworthiness Rating
1	FLO Labeling International (Flo-Cert)	A
2	Dave & Nic Certification	C
3	Non Violent Dove Certification	D

6. CIDIBB IN PRACTICE

In order to better illustrate potential uses of CIDIBB, this section provides two potential scenarios of its use in practice. These scenarios depict the use of CIDIBB by Lucy (Consumer Advocate) and William (Certifier).

6.1. Scenario #1: A Consumer Advocate Uses CIDIBB to Create a New Product or Service Rating System

Lucy is the CEO of a well-established product rating firm. Lucy's firm is an information aggregator that harvests information about sustainable consumer products and publishes proprietary product ratings (organized by UPC code). The firm has created a number of apps that allow consumers to access the product ratings while they are shopping either in a physical store or online. Their business model is to sell a low cost-subscription of their service to individual consumers. One of the early entrants into this market niche was GoodGuide.

In Lucy's business model, consumer values are expressed as concerns and questions, which can be translated into machine-understandable queries. These queries are executed against standardized data and semantically enriched by CIDIBB ontologies. For example, some consumers may be concerned if child labor was used during production processes. This concern can be translated into a machine-understandable query as presented below:

```
If <NoEmployementOfChildrenUnderAgeOf15 hasEvaluationDecision some
'Evaluation Decision'> and
<NoEmployementOfChildrenUnderAgeOf15 hasCriteriaType value "Core
Criteria"> and
<NoEmployementOfChildrenUnderAgeOf15 hasTimeline value "Initial
Audit"> and
<NoEmployementOfChildrenUnderAgeOf15 hasApplicability value
"Members of Organization"> and
<NoEmployementOfChildrenUnderAgeOf15 hasIndicator value "There are
no Children under the age of 15 years employed">
```

If the returned query result is true, then it means that no child labor was used in the production of the good. Query results are then fed into the rating algorithm. The output of the algorithm is one or more ratings that reflect the value of the good or service according to a particular value system. The automatic nature of the process allows Lucy to target as many specific consumer values and interests as she wants.

Moreover, Lucy also has a tool to provide customizable assessment of the trustworthiness of the information. Lucy can provide the consumer with the 28 use case questions as a default set of questions to define the trustworthiness of a certification. Based on the consumer's own interpretation of trustworthiness of a certification, the consumer can choose either to directly use or edit this default set of questions, deselecting some of these 28 questions, adding more questions, and assigning different weights to these questions using functions available in Lucy's apps. The finalized set of questions that defines the consumer's understanding of trustworthiness of a certification will be translated into machine-understandable queries, and then the queries will be executed against data structured and organized with CIDIBB ontologies. The query results will be a sequence of true or false answers to the set of questions that defines the consumer's understanding of trustworthiness of a certification. A trustworthiness score of the certification can be calculated based on the sequence of answers and weights assigned to these questions. The sequence of true or false answers, the trustworthiness score, and the algorithm used to calculate the score can all be presented to the consumer as the answer to his/her query regarding the extent to which a certification associated with a particular product is trustworthy.

6.2. Scenario #2: A Certifying Organization Uses CIDIBB to Create a New Virtual Certificate

William is the co-founder of Cyber-Just Trade (CJT), a start-up certification agency. He envisioned the creation of first-ever virtual sustainable certification scheme as the company's lever to compete against other much bigger certification agencies. William and his co-founder soon confronted with three major challenges to their efforts: a) the ownership of certification information is in the hands of the applicant and not the certification agency, b) commercial privacy related to certification data for each firm in a supply chain, and c) provision of instant traceability and comparability requires the availability of standardized data across supply chain firms and other certification schemes. Upon discovering CIDIBB, William realizes it can help his certification agency in overcoming the above-mentioned challenges.

William can use the CIDIBB taxonomy to create standardization of all the certification inspection reports. The standardization could facilitate instant traceability and comparability. William creates the certification report in the CIDIBB format by tagging each of their reporting elements with standardized concepts in the CIDIBB taxonomy. For instance, William certification reports can be seen as below:

```
<hasIdentifier>FLO_1341</hasIdentifier>
<Applicant hasIdentifier="FL01341">Carlos</Applicant>
<hasLocation>Mexico</hasLocation>
<hasCertificationOfficer>Janssen Martina</hasCertificationOfficer>
```

William could also use the 28 trustworthiness question from the CIDIBB framework to conduct internal verification of his certification and assess how his certification ranks compare to his competitors. Since the CIDIBB is based on three ontologies, William could adjust the level of abstraction in reporting the certification results depending to the consent from the producer, data owner of certification. William could use the CIDIBB's three ontologies in appealing to the data owner regarding the disclosure of the certification results. At the highest level of abstraction, the reporting could be based on the CertIN taxonomy and at the highest level of details to use both the CertIN and the local ontology.

CONCLUSION

Global markets for information-intensive products contain sharp information asymmetries that lead to public problems such as market inefficiencies, resulting in consumer purchasing decisions that are based on incomplete information. Unintended side effects of these information asymmetries vary depending on the markets in question, ranging from negative externalities such as environmental degradation in the case of unsustainable production practices for agricultural products, loss of human capital in the case of exploitative labor practices, or unfavorable patient outcomes in the case of incomplete information about the quality of care provided in different health care settings or the addictiveness of opioid pain medications such as Oxycodone. Elimination or reduction of such information asymmetries has long been the goal of governments as well as various non-governmental entities that recognize that addressing issues such as sustainable production, socially just labor practices and reduction in energy needs and health expenditure is closely linked to consumers who are fully aware of the economic, environmental and social impacts of their purchasing decisions.

The current research explored creation of ontology-enabled interoperable data infrastructure, based on the semantic web that would enable information sharing and collaboration in traditionally information-restricted markets. Throughout the three-year project, the feasibility of tagging and broadcasting a diverse and dispersed set of data from product certification and inspection processes to allow for assessment of their accuracy and trustworthiness was explored. The main technical result of this project is a proof-of-concept Certification and Inspection Data Infrastructure Building Block

(CIDIBB), which is a set of data standards built on semantic web applications and the functionalities of a formal ontology of the certification and inspection process. While the current proof-of-concept focuses narrowly on certified fair-trade coffee and its functionality is limited, it has the potential to become universally applicable to any certification and inspection process for any product and service. The evaluation of CIDIBB presented in this paper suggests that governance mechanisms to make data available are indispensable in the certification domain, and even in the ideal benchmark, 5 out of 28 questions require governance intervention to be answered, which translates in necessary public-private collaborations to ensure the existence of such governance mechanisms.

CIDIBB has a direct impact on the certification industry by providing a tool to promote transparency and trust in certification systems, leading to the possibility of having virtual certificates that provide detailed information about certification and inspection processes as a measure of trust of a given certificate. Moreover, CIDIBB will also have broader impacts in a diversity of domains and industries using certification schemes. CIDIBB constitutes a platform to be used to promote ethical consumption; helping consumers to make better decisions on the basis of trusted information about product characteristics. Additionally, CIDIBB may also contribute by providing a platform for those interested in sustainability to find trusted partners and providers to produce goods and deliver services in more sustainable ways. More generally, a CIDIBB is a key component for any kind of data sharing initiative given that it provides a tool to assess the trustworthiness of the data being shared in any domain.

Achieving universal applicability of the CIDIBB, however, requires a series of steps aimed at refinement and broadening of this existing proof-of-concept and gradually increasing the scope of products and services. The first step is to further refine and test a full prototype in the original area of its focus, namely certified fair-trade coffee. Such refinement and testing requires access to real world certification and inspection data. The second step is the application of the refined CIDIBB to other certifications surrounding coffee, such as organic, to test the applicability of Certification and Inspection Ontology (CertIN) to other certification schemes. The continual focus on coffee takes advantage of our domain expertise and allows us to test CIDIBB's ability to address comparability of different certification schemes. If such buildup is successful, the next step toward testing for universal applicability is to incorporate other agricultural products that might require different inspection processes. Finally, the last step toward universal application is to use the existing CIDIBB for non-agricultural domains.

Making CIDIBB a reality requires integration of data and information that are under the ownership and stewardship of public and private entities. In this way, many non-technical requirements also need to be met. While information quality and integrity have always been an issue of concern even in situations with a single information source, it will be an even more complex problem in the case of a platform that is designed to integrate information from multiple disparate sources. Thus, creating technical and process mechanisms to ensure information integrity and security is essential for the data to be trustworthy. Moreover, designing information policy that balances the need for supply chain transparency and ability of businesses to remain competitive is key. Establishing a governance structure is crucial for all large system development projects, but perhaps especially so for the development of platforms dealing with the complex determinants of sustainability such as CIDIBB. The key to this process is establishing a basis for "principled engagement,"—a common understanding of the ways in which different stakeholders use central concepts and terms (Emerson, Nabatchi & Balogh, 2012).

By making the proof-of-concept CIDBB operational, it will provide, for the first time, a way for end users to reduce sharp information asymmetries in consumer markets through access to certification and inspection information in areas as widely dispersed as the performance of a health care provider, energy consumption patterns, or the safety of products we use each day in our daily routines.

In the current version, on the other hand, findings focused predominantly on the application area of product information including labeling and certification specifically using the case of fair trade coffee certification. This paper primarily focuses the vignette on two users of CIDIBB, namely: consumer's advocate (information aggregator) and certifier / certification body. However, we argue that CIDIBB has the potential for becoming universally applicable to any certification and inspection process for any product and service. Nonetheless, we also realize that to extend the applicability of CIDIBB, future research is needed. This paper also focuses more on explaining the work and environment of CIDIBB, and less on providing detail technical description of each process within CIDIBB. For instance, we did not describe the process to recast standard tables of data in the CID as a RDF triple store. As one final point worth considering, future research could consider using the design science approach (see Hevner et al., 2004 or Peffers et al., 2007) for presenting and providing insight of the development processes of information system artifacts.

ACKNOWLEDGMENT

This research is partially supported by the National Science Foundation under Grants No. 37656 and IIS-0540069, the Consejo Nacional de Ciencia y Tecnología under Grants No. 84082 and 133670. Any opinions expressed in this material are those of the authors and do not necessarily reflect the views of NSF or CONACYT. Partial support provided for JL by Rensselaer Polytechnic Institute and Indiana University.

REFERENCES

Albersmeier, F., Schulze, H., Jahn, G., & Spiller, A. (2009). The reliability of third-party certification in the food chain: From checklists to risk-oriented auditing. *Food Control*, *20*(10), 927–935. doi:10.1016/j.foodcont.2009.01.010

Arora, R. (2006). Product positioning based on search, experience and credence attributes using conjoint analysis. *Journal of Product and Brand Management*, *15*(4/5), 285–292. doi:10.1108/10610420610685695

Berners-Lee, T. (2006). Linked Data. Retrieved from http://www.w3.org/DesignIssues/LinkedData.html

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, *284*(5), 34–43. doi:10.1038/scientificamerican0501-34 PMID:11396337

Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked Data—The Story So Far. *International Journal on Semantic Web and Information Systems*, *5*(3), 1–22. doi:10.4018/jswis.2009081901

Campbell, H., Murcott, A., & MacKenzie, A. (2010). Kosher in New York City, halal in Aquitaine: Challenging the relationship between neoliberalism and food auditing. *Agriculture and Human Values*, *28*(1), 67–79. doi:10.1007/s10460-010-9260-3

Colquitt, J. A., Scott, B. A., & LePine, J. A. (2007). Trust, trustworthiness, and trust propensity: A meta-analytic test of their unique relationships with risk taking and job performance. *The Journal of Applied Psychology*, *92*(4), 909–927. doi:10.1037/0021-9010.92.4.909 PMID:17638454

Davenport, C., & Hakim, D. (2016, January 4). U.S. Sues Volkswagen in Diesel Emissions Scandal. *The New York Times*. Retrieved from http://www.nytimes.com/2016/01/05/business/vw-sued-justice-department-emissions-scandal.html

Deaton, B. J. (2004). A theoretical framework for examining the role of third-party certifiers. *Food Control*, *15*(8), 615–619. doi:10.1016/j.foodcont.2003.09.007

Emerson, K., Nabatchi, T., & Balogh, S. (2012). An integrative framework for collaborative governance. *Journal of Public Administration: Research and Theory*, *22*(1), 1–29. doi:10.1093/jopart/mur011

Fox, P., & McGuinness, D. L. (2008). TWC Semantic Web Methodology. Retrieved from http://tw.rpi.edu/web/doc/TWC_SemanticWebMethodology

Graham, M., & Haarstad, H. (2011). Transparency and Development: Ethical Consumption Through Web 2.0 and the Internet of Things. *Information Technologies and International Development*, *7*(1).

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199–220. doi:10.1006/knac.1993.1008

Gruninger, M., & Fox, M. S. (1995). Methodology for the Design and Evaluation of Ontologies. In *Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing IJCAI '95*, Montreal.

Hart, P., & Saunders, C. (1997). Power and Trust: Critical Factors in the Adoption and Use of Electronic Data Interchange. *Organization Science*, *8*(1), 23–42. doi:10.1287/orsc.8.1.23

Henderson, D., Sheetz, S. D., & Trinklec, B. S. (2012). The determinants of inter-organizational and internal in-house adoption of XBRL: A structural equation model. *International Journal of Accounting Information Systems*, *13*(2), 109–140. doi:10.1016/j.accinf.2012.02.001

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *Management Information Systems Quarterly*, *28*(1), 75–105. doi:10.2307/25148625

Jahn, G., Schramm, M., & Spiller, A. (2005). The reliability of certification: Quality labels as a consumer policy tool. *Journal of Consumer Policy*, *28*(1), 53–73. doi:10.1007/s10603-004-7298-6

Jarman, H., & Luna-Reyes, L. F. (Eds.). (2016). *Private Data and Public Value: Governance, Green Consumption, and Sustainable Supply Chains. Public Administration and Information Technology (PAIT) Series*. Springer. doi:10.1007/978-3-319-27823-0

LandXML. (2017). About LandXML. Retrieved from http://www.landxml.org/About.aspx

Luna-Reyes, L. F., Zhang, J., Roy, R., Andersen, D. F., Whitmore, A., & Andersen, D. L. (2013). Information strategies to support full information product pricing: The role of trust. *Information Polity: The International Journal of Government & Democracy in the Information Age*, *18*(1), 75–91.

Luna-Reyes, L. F., Zhang, J., Whitmore, A., Jarman, H., Picazo-Vela, S., Andersen, D. L., & Andersen, D. F. et al. (2014). Full information product pricing: An information strategy for harnessing consumer choice to create a more sustainable world. *Communications of the Association for Information Systems*, *34*(1), 637–654.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model Of Organizational Trust. *Academy of Management Review*, *20*(3), 709–734. doi:10.5465/AMR.1995.9508080335

McKnight, D. H., & Chervany, N. L. (2001). Trust and Distrust Definitions: One Bite at a Time. In R. Falcone, M. Singh, & Y.-H. Tan (Eds.), *Trust in Cyber-societies* (pp. 27–54). Springer Berlin Heidelberg; Retrieved from http://link.springer.com/chapter/10.1007/3-540-45547-7_3 doi:10.1007/3-540-45547-7_3

Mitchell, S. (2011). Walmart's greenwash: Why the retail giant is still unsustainable. Retrieved from https://grist.org/series/2011-11-07-walmart-greenwash-retail-giant-still-unsustainable/

Nidumolu, S. R. (1989). The Impact of Interorganizational Systems on the Form and Climate of Seller-Buyer Relationships: A Structural Equations Modelling Approach. In *Proceedings of the 10th International Conference on Information Systems* (pp. 289-304). doi:10.1145/75034.75059

Noy, N. (2005). Ontology mapping and alignment. In Fifth International Workshop on Ontology Matching collocated with the 9th International Semantic Web Conference ISWC-2010, Shangai, China. Retrieved from http://andrei.clubcisco.ro/cursuri/5master/aac-sac/misc/Ontology%20Mapping%20and%20Alignment.pdf

Pavlou, P. A., & Gefen, D. (2004). Building Effective Online Marketplaces with Institution-Based Trust. *Information Systems Research*, *15*(1), 37–59. doi:10.1287/isre.1040.0015

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77. doi:10.2753/MIS0742-1222240302

Powell, W. (1996). Trust-Based Forms of Governance. In T. Tyler & R. M. Kramer (Eds.), *Trust in organizations: Frontiers of theory and research* (pp. 51–67). Thousand Oaks, CA: Sage Publications. doi:10.4135/9781452243610.n4

Puranam, P., & Vanneste, B. S. (2009). Trust and Governance: Untangling A Tangled Web. *Academy of Management Review*, *34*(1), 11–31. doi:10.5465/AMR.2009.35713271

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, *23*(3), 393–404. doi:10.5465/AMR.1998.926617

Sayogo, D. S., Ran, W., Tayi, G. K., Lucinano, J. S., Luna-Reyes, L. F., Depaula, N., & Pardo, T. A. et al. (2016). Ontological Modeling of Certification and Inspection Process to Support Smart Disclosure of Product Information. *International Journal of Public Administration in the Digital Age*, *3*(2), 86–108. doi:10.4018/IJPADA.2016040106

Starobin, S., & Weinthal, E. (2010). The Search for Credible Information in Social and Environmental Global Governance: The Kosher Label. *Business and Politics*, *12*(3), 1–35. doi:10.2202/1469-3569.1322

Stigler, G. J. (1957). Perfect Competition, Historically Contemplated. *Journal of Political Economy*, *65*(1), 1–17. doi:10.1086/257878

Tanner, B. (2000). Independent assessment by third-party certification bodies. *Food Control*, *11*(5), 415–417. doi:10.1016/S0956-7135(99)00055-9

Uschold, M., & Gruninger, M. (1996). Ontologies: Principles, methods and applications. *The Knowledge Engineering Review*, *11*(2), 93–136. doi:10.1017/S0269888900007797

Wang, W., & Benbasat, I. (2005). Trust in and adoption of online recommendation agents. *Journal of the Association for Information Systems*, *6*(3), 4.

Williamson, O. E. (1993). Calculativeness, Trust, and Economic Organization. *The Journal of Law & Economics*, *36*(1), 453–486. doi:10.1086/467284

Zhang, J., Liu, H., Sayogo, D. S., Picazo-Vela, S., & Luna-Reyes, L. F. (2016). Strengthening institutional-based trust for sustainable consumption: Lessons for smart disclosure. *Government Information Quarterly*, *33*(3), 552–561. doi:10.1016/j.giq.2016.01.009

ENDNOTES

1 See for example, FSC (Forest Stewardship Council) https://us.fsc.org/en-us/certification/chain-of-custody-certification, and UTZ Certified Chain of Custody https://utzcertified.org/ndp?article&id=26584817

2 The name "data infrastructure building block" derives from the National Science Foundation Data Infrastructure Building Block program which aims to "foster cross-community infrastructure development that solves common problems, while building blocks of data infrastructure that can support and provide data solutions to a broader range of scientific disciplines while reducing duplicative efforts." (http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504776). A data infrastructure promotes data sharing and consumption through a common data structure and standard.

3 For a brief description on greenwashing, see (Mitchell, 2011)

Joanne S. Luciano is a Visiting Associate Professor in the School of Informatics, Computing, and Engineering, Indiana University Bloomington, where she teaches graduate courses in Data Science. Dr. Luciano is an internationally recognized multidisciplinary scientist and consulting engineer leading research, development, technology, and education. She works at the intersection of the life sciences, mathematics, and knowledge technology with degrees in computer science and in cognitive and neural systems. She has over 25 years in technology research and development, product development, consulting, and business development in organizations ranging from start-ups to large corporations. Dr. Luciano's professional roles have included project management, quality engineering and assurance, knowledge representation, intellectual property assessment, competitive analysis, patent infringement, due diligence, forensic research, and business development strategy. Current research interests include Health Web Science, Predictive, Preventive, Personalized and Participatory (P4) Medicine, Semantic Web technologies, ontologies, knowledge representation, wearable devices, and Internet of Things. Dr. Luciano authored a recent book, "Health Web Science," and holds two US patents in predictive medicine.

Djoko Sigit Sayogo is an assistant professor in the Department of Economics and Business at the University of Muhammadiyah at Malang (UMM), Indonesia. Additionally, he was the Vice Director of Research at UMM. His research interests focus on the intersection of information technology, public policy, management, and economic development and on the broad topic of the application of information technology both in public and private entities, including subjects such as information sharing in the financial market, open data, smart disclosure policies, e-financial reporting, and open budget.

Weijia Ran is a doctoral graduate from the Department of Informatics at the University at Albany, State University of New York. Her research interests include e-learning in the workplace, knowledge management, information systems in organizations, complex systems, and sustainability.

Nic DePaula is a Ph.D. candidate in the Department of Information Science and lecturer in the departments of Information Technology Management and Information Science at the University at Albany, State University of New York. He conducts research in the areas of e-government, information policy, and new media.

Holly Jarman is the John G. Searle Assistant Professor within the Department of Health Management and Policy at the University of Michigan. She researches the impacts of regulatory politics on health in cross border contexts.

Giri Kumar Tayi is a Professor of Management Science and Information Systems at the State University of New York at Albany. He obtained his Ph.D. from Carnegie Mellon University and his research and teaching interests are interdisciplinary and span the fields of Information Systems, Operations Management and Operations Research. He has published about 60 refereed journal article. Many of the articles appear in top-tier academic journals. He has Co-guest edited nine special journal issues. He serves or has served on the Editorial Boards of academic journals such as Information Systems Research (AE and SE), POMS (SE), IEEE Intelligent Systems, IEEE Computing Now, Decision Sciences Journal, ACM Journal of Data and Information Quality, Information Technology Management, Information Systems Frontiers, Information and Management, International Journal of Shipping and Transport Logistics.

Jing Zhang is a Professor of Information Management and Business Analytics at Clark University. Her research focuses on inter-organizational information and knowledge sharing, organizational impact of technology and innovation, and the information technology support for sustainable consumption and sustainable supply chain. She served as the program co-chair for dg.o conference from 2012-2015, and she is an elected board member of the Digital Government Society. Her publications appear in European Journal of Information Systems, Government Information Quarterly, Public Performance and Management Review, Information Technology and Management, among others.

Jana Hrdinová is the Associate Director at the Ohio State Center for Real Estate. Prior to coming to the Center Jana was a Program Associate at Center for Technology in Government at the University at Albany, SUNY, where she was responsible for project development, planning, implementation, management, analysis, and writing. Prior to entering the academic world, Jana worked for U.S. Senator (Minnesota) Mark Dayton as deputy director for constituent services concentrating predominantly on immigration issues. In addition to assisting constituents with problems regarding federal agencies, Jana conducted outreach to Minnesota's African-immigrant and African-American communities, and advised Senator Dayton on immigration, worker's compensation, and postal service policy issues. Jana holds a Master of Arts in Political Science and a Bachelor of Arts in International Relations from the University of Minnesota.

Theresa A. Pardo is the Director of the Center for Technology in Government at the University at Albany, State University of New York, where she is also a Full Research Professor in Public Administration and Policy. Dr. Pardo fills a number of leadership and advisory roles including serving as OpenNY Adviser to New York State's Governor Cuomo, as a member of the U.S. Environmental Protection Agency's National Advisory Committee and as a member of the Editorial Board of Government Information Quarterly. Dr. Pardo is founder of the Smart Cities, Smart Government Research-Practice Consortium and is a Past-President of the Digital Government Society.

Deborah Lines Andersen is professor emeritus of information science and policy at the University at Albany. Her areas of specialization are information policy, research design and library/information science. She retired in 2016.

David Andersen is the O'Leary Distinguished Service Professor of Public Administration, Public Policy, and Information Science at the Rockefeller College, University at Albany. His work centers on applying system dynamics, systems thinking, and information technology approaches to problems in the public, not-for-profit, and private sectors.

Luis Felipe Luna-Reyes is an Associate Professor in the Department of Public Administration and Policy, Affiliated Faculty in the Department of Information Science at the University at Albany. He has been a Fulbright Scholar and he is currently Faculty Fellow at the Center for Technology in Government. He is also a Research Affiliated at the Universidad de las Americas, Puebla and a member of the Mexican National Research System. His research is at the intersection of Public Administration, Information Systems and Systems Sciences. More specifically, Luna-Reyes use multi-method approaches to contribute to a better understanding of collaboration and governance processes in the development of information technologies across functional and organizational boundaries in government. His research interests are related to areas such as inter-organizational collaboration, information sharing, success of government-wide Web sites, modeling and simulation, and information policy. He is the author or co-author of more than 100 articles published in leading Journals and Academic Conferences such as Government Information Quarterly, European Journal of Information Systems, System Dynamics Review, Information Polity, and Information Technology and Management among others, working with a network of more than 60 co-authors around the world. His research has been funded by the Consejo Nacional de Ciencia y Tecnología in México, the National Science Foundation, the Fondo de Información y Documentación para la Industria (Infotec) in México, and the Economic Commission for latin America and the Caribbean (ECLAC).

Building a Certification and Inspection Data Infrastructure to Promote Transparent Markets

ORIGINALITY REPORT

6%	3%	9%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	"Collaboration and Trust Building Among Public and Private Actors", Private Data and Public Value, 2016. Publication	2%
2	Taewoo Nam. "Achievable or Ambitious?: A Comparative and Critical View of Government 3.0 in Korea", International Journal of Electronic Government Research, 2017 Publication	1%
3	"Public Value and Private Organizations", Private Data and Public Value, 2016. Publication	1%
4	Xiaogang Ma, Stace E. Beaulieu, Linyun Fu, Peter Fox, Massimo Di Stefano, Patrick West. "chapter 5 Documenting Provenance for Reproducible Marine Ecosystem Assessment in Open Science", IGI Global, 2017 Publication	1%
5	Quan Chen, Jiangtao Wang, Ruiqiu Ou, Sang-Bing Tsai. "An Empirical Study on the	1%

Network Model and the Online Knowledge Production Structure", Journal of Information Technology Research, 2019

Publication

6	d-nb.info Internet Source	1%

Exclude quotes	On	Exclude matches	< 1%
Exclude bibliography	On		